CONTENTS

7 Interval Estimation

7.1 Confidence Intervals

Definition 1. Confidence Interval
An interval

$$(l(x_1, \ldots, x_n), u(x_1, \ldots, xn))$$

is called a $100\gamma\%$ confidence interval for θ if

$$P[l(x_1, \ldots, x_n) < \theta < u(x_1, \ldots, xn)] = \gamma$$

where $0 < \gamma < 1$.

The observed values $l(x_1, \ldots, x_n)$ and $u(x_1, \ldots, x_n)$ are called lower and upper confidence limits, respectively.

Definition 2. One-Sided Confidence Limits If

$$P[l(x_1, \ldots, x_n) < \theta] = \gamma$$

then $l(x_1, \ldots, x_n)$ is called a one-sided lower $100\gamma\%$ confidence limit for θ.

If

$$P[u(x_1, \ldots, x_n) > \theta] = \gamma$$

then $u(x_1, \ldots, x_n)$ is called a one-sided upper $100\gamma\%$ confidence limit for θ.

In general, if (θ_L, θ_U) is a $100\gamma\%$ confidence interval for a parameter θ, and if $\tau(\theta)$ is a monotonic increasing function of $\theta \in \Omega$, The $(\tau(\theta_L), \tau(\theta_U))$ is a $100\gamma\%$ confidence interval for $\tau(\theta)$.

Example 1. Consider a random sample of size n from an exponential distribution, $X_i \sim Exp(\theta)$.

(a) Construct a one-sided lower $100\gamma\%$ confidence limit for θ.

(b) Construct a one-sided upper $100\gamma\%$ confidence limit for θ.

(c) Construct a $100\gamma\%$ confidence interval for θ.

(d) Find a one-sided lower $100\gamma\%$ confidence limit for $P(X > t) = e^{-t/\theta}$.

Example 2.

Consider independent random samples from two gamma distributions, $X \sim gamma(4, \beta_1)$ and $Y_j \sim gamma(10, \beta_2)$; $i = 1, \ldots, n_1$, $j = 1, \ldots, n_2$.

(a) Find the distribution of $\left(\frac{\beta_2}{\beta_1}\right)\left(\frac{5\bar{X}}{2\bar{Y}}\right)$.

(b) Derive a $100(1-\alpha)\%$ confidence for $\frac{\beta_2}{\beta_1}$.

Example 3.

Consider a random sample of size 40 from a uniform distribution, $X_i \sim U(0, \theta)$, $\theta > 0$, and let $X_{n:n}$ be the largest order statistic. Find the constant c such that $(x_{n:n}, cx_{n:n})$ is a 92% confidence interval for θ.

7.2 Pivotal Quantity Method

Definition 3. Pivotal Quantity
If $Q = q(X_1, \ldots, X_n; \theta)$ is a random variable that is a function only of $(X_1, \ldots, X_n)$ and θ, then Q is called a pivotal quantity if its distribution does not depend on θ or any other unknown parameters. That is, if $X \sim F(\mathbf{x}|\theta)$, then Q has the same distribution for all values of θ.

Example 4. (Gamma pivot)
Suppose that $X_1, \ldots, X_n$ are iid $Exp(\theta)$, find the pivotal quantity based on the sufficient statistics $T = \sum X_i$.

Example 5.
Consider a random sample from a normal distribution, $X \sim N(\mu, \sigma^2)$, where both μ and σ^2 are unknown. If $\hat{\mu}$ and $\hat{\sigma}$ are the MLEs of μ and σ,

(a) show that $\frac{\hat{\mu}-\mu}{\hat{\sigma}}$ and $\hat{\sigma}/\sigma$ are pivotal quantities;

(b) find a $100(1-\alpha)\%$ confidence interval for μ.

(c) find an equal tail $100(1-\alpha)\%$ confidence interval for σ^2.

Example 6.

Let $X_1, X_2, \ldots, X_n$ be a random sample from a Weibull distribution, $X \sim WEI(\theta, 4)$.

(a) Show that $Q = 2\sum_{i=1}^{n} X_i^4/\theta^4 \sim \chi^2(2n)$.

(b) Use Q to derive an equal tailed $100\gamma\%$ confidence interval for θ.

Example 7.

Let $X_1, \ldots, X_n$, be a random sample from a gamma distribution with parameters $\alpha = 5$ and unknown θ.

(a) Find a pivotal quantity for the parameter θ based on the sufficient statistic.

(b) Derive an equal tail 92% confidence interval for θ based on the pivotal qualtity from part (a).

It may not always be possible to find a pivotal quatity, but for a sample from a continuous distribution with a single unknown parameter, at least one pivotal quantity can always be derived by use of the probability integral transform.
If

$$X \sim f(x;\theta)$$

and if

$$F(x;\theta)$$

is the CDF of X, then

$$F(X;\theta) \sim U(0,1)$$

and consequently

$$Y_i \sim -\ln F(X_i,\theta) \sim EXP(1).$$

For a random sample $X_1,\ldots,X_n$, it follows that

$$-2\sum_{i=1}^{n} \ln F(X_i;\theta) \sim \chi^2(2n)$$

so that

$$P[\chi^2_{\alpha/2}(2n) < -2\ln F(X_i;\theta) < \chi^2_{1-\alpha/1}(2n)] = 1-\alpha$$

and inverting this statement will provide a confidence region for θ.
If the CDF is not in closed form or if it is too complicated, then the inversion may have to be done numerically.
If $F(x;\theta)$ is a monotonic increasing (or decreasing) function of θ, then the resulting confidence region will be an interval.

Notice also that $1 - F(X_i;\theta) \sim U(0,1)$ and

$$-2\sum_{i=1}^{n} \ln[1 - F(X_i;\theta)] \sim \chi^2(2n)$$

Example 8.

Consider a random sample from a Pareto distribution, $X \sim PAR(\alpha, \theta = 300)$, find a $100(1 - \alpha)\%$ confidence interval for α.

7.3 Aproximate Confidence Intervals

For discrete distributions, and for some multiparameter problems, a pivotal quantity may not exist. However, an approximate pivotal quantity often can be obtained based on asymptotic resultys. Let $X_1, \ldots, X_n$ be a random sample from a distribution with pdf $f(x; \theta)$. As noted in previous chapter, MLEs are asymptotically normal under certain condition.

Example 9.
Consider a random sample from a Bernoulli distribution, $X \sim BIN(1, p)$. Find an approximate confidence limits for p.

7.4 Credible Interval

A credible interval (or in general, a credible set) is the Bayesian analogue of a confidence interval. A $100(1-\alpha)\%$ credible set C is a subset of Θ such that

$$\int_C \pi(\theta|\mathbf{x})d\theta = 1 - \alpha$$

If the parameter space Θ is discrete, a sum replaces the integral.

Definition 4.
If a is the $\frac{\alpha}{2}$ posterior quantile for θ, and b is the $1-\frac{\alpha}{2}$ posterior quantile for θ, then (a, b) is a $100(1-\alpha)\%$ **equal probability credible interval** for θ.

Example 10.

The following amounts were paid on a hospital liability policy:

$$125, 138, 142, 103, 137, 311, 127, 102, 144, 231.$$

The amount of a single payment has the single-parameter Pareto distribution with $\theta = 102$ and α unknown. The prior distribution has the gamma distribution with $\alpha = 3$ and $\theta = 1$. Determine the 96% equal probability credible interval for α.

The equal-tail credible interval approach is ideal when the posterior distribution is symmetric. If $\pi(\theta|\mathbf{x})$ is skewed, a better approach is to create an interval of $\theta-$values having the Highest Posterior Density (HPD).

Definition 5.

A $100(1-\alpha)\%$ HPD region for θ is a subset $C \in \Theta$ defined by

$$C = \{\theta : \pi(\theta|\mathbf{x}) \geq k\}$$

where k is the largest number such that

$$\int_{\theta:\pi(\theta|\mathbf{x})\geq k} \pi(\theta|\mathbf{x})d\theta = 1 - \alpha$$

The value k can be thought of as a horizontal line placed over the posterior density whose intersection(s) with the posterior define regions with probability $1 - \alpha$.

Theorem 1.

If the posterior random variable $\theta|\mathbf{x}$ is continuous and unimodal, then the $100(1-\alpha)\%$ HPD credible interval is the unique solution to

$$\int_a^b \pi(\theta|\mathbf{x})d\theta = 1-\alpha$$

$$\pi(a|\mathbf{x}) = \pi(b|\mathbf{x})$$

Example 11.

You are given the following:

$$f(x|\theta) = \frac{5x^4}{\theta^5}, 0 < x < \theta.$$

$$\pi(\theta) = \frac{6}{\theta^7}, \theta > 1.$$

Three observations were observed: 500, 600, 1300. Find a 95% "HPD" credible set for θ.

Example 12.

The following amounts were paid on a hospital liability policy:

125	132	141	107	133
319	126	104	145	223

The amount of a single payment has the single-parameter Pareto distribution with $\theta = 100$ and α unknown. The prior distribution has the gamma distribution with $\alpha = 2$ and $\theta = 1$. Determine the 95% HPD credible interval for α.

$\boxed{a=1.1832,\ b = 3.9384}$


```
f = function(x){
y = numeric(2)
y[1] = pgamma(x[2],12,4.801121) - pgamma(x[1],12,4.801121) - 0.95
y[2] = dgamma(x[1],12,4.801121) - dgamma(x[2],12,4.801121)
y
}
library(nleqslv)
xstart = c(1,3)
nleqslv(xstart, f, control=list(btol=.01),
method="Newton")
```

Example 13.

Let $X_1, \ldots, X_n$, be a random sample from $N(\theta, 1)$. Assume that the prior distribution of Θ is $N(\mu, \sigma^2)$ with known μ and σ^2.

(a) Derive the posterior distrbution of Θ.

(b) Find a Bayesian interval of θ with confidence coefficient $1 - \alpha$.

(c) Find the corresponding non-Bayesian confidence interval of θ using pivotal quantity method.